DOCUMENT RESUME

ED 169 893 IR 007 134

TITLE	Data Systems of the National Center for Health Statistics.
INSTITUTION	National Center for Health Statistics (DHEW), Rockville, Md.
REPORT NO	DHEW-PHS-79-1247
PUB DATE	Jan 79
NOTE	55p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (622-857/615)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Data Collection; Data Processing; Health Services; *Information Dissemination; *Information Services; Information Sources; Information Utilization; *National Surveys; Physical Health; *Public Health; Questionnaires; Statistical Data

ABSTRACT

This brief, semi-technical handbook describes the various data systems of the National Center for Health Statistics (NCHS) with summaries of background, purpose and scope, sample design, data collection procedures, questionnaire content, data release, and data uses for individual NCHS surveys. NCHS operates a diverse survey and inventory program that encompasses its legislative authorization to collect statistics on (1) the extent and nature of illness and disability of the population of the United States; (2) the impact of illness and disability of the population on the economy of the United States and on other aspects of the well-being of its population; (3) health resources; (4) utilization of health care; (5) health care costs and financing; and (6) family formation, growth, and dissolution. This summary is intended to serve the needs of visitors to the NCHS, as well as those of others at the federal, state, and local levels who need to understand the structure and scope of its programs. (Author/CWM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED169893

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

data systems

of the National Center for Health Statistics

DHEW Publication No.
(PHS) 79-1247

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Public Health Service
Office of the Assistant Secretary for Health
National Center for Health Statistics

Hyattsville, Maryland
January 1979

IR007134

Foreword

This brief, semi-technical handbook describing the various data systems of the National Center for Health Statistics has been created to serve a variety of needs. The greatest need was for a publication bringing together descriptive materials under one cover. For the user who needs more detailed, technical knowledge of these systems, this is available in separate publications for the individual programs. It is expected that this summary publication will also serve the needs of many visitors to the Center, as well as those of others at the Federal, State, and local levels needing to understand the structure and scope of its programs.

One feature that should be carefully noted is the flexible nature of many of these surveys. They, while aware of the need for standard measurement over time, also change with the demands for new data. Systematic reviews lead to careful revisions and additions as the needs are identified. When necessary, new survey activities are also designed and fielded. Thus, periodic changes in this handbook are anticipated as these revisions are made.

Elijah L. White
Associate Director for Data Systems

TABLE OF CONTENTS

Introduction---------- 1

Vital Statistics---------- 5

Basic Vital Statistics---------- 5

Vital Statistics Followback Surveys---------- 9

National Survey of Family Growth---------- 13

General Population Surveys---------- 17

Health Interview Survey---------- 17

Health and Nutrition Examination Survey---------- 23

Health Resources Utilization Statistics---------- 28

Hospital Discharge Survey---------- 28

National Ambulatory Medical Care Survey---------- 31

National Nursing Home Survey---------- 36

National Reporting System for Family Planning Services-------- 40

Health Resources Statistics---------- 44

Health Facilities Data Systems---------- 44

Master Facility Inventory---------- 44

National Inventory of Family Planning Clinics---------- 47

Health Manpower Inventories and Surveys---------- 50

INTRODUCTION

The National Center for Health Statistics (NCHS) is one of the major Federal statistical organizations. NCHS operates a diverse survey and inventory program that encompasses its legislative authorization to collect statistics on: (a) the extent and nature of illness and disability of the population of the United States, including life expectancy, the incidence of various acute and chronic illnesses, and infant and maternal morbidity and mortality; (b) the impact of illness and disability of the population on the economy of the United States and on other aspects of the well-being of its population; (c) health resources, including health professionals by specialty and type of practice, and the supply of services by hospitals, extended care facilities, home health agencies, and other health institutions; (d) utilization of health care, including utilization of ambulatory health services, services of hospitals, extended care facilities, home health agencies and other institutions; (e) health care costs and financing; and (f) family formation, growth, and dissolution.

This document provides brief descriptions of each of the data systems operated by NCHS. Each project description contains a statement of the purpose of the data system, the contents of the data collection instrument, an overview of the data collection process, and an indication of the availability of the data to the public.

Confidentiality of Data

Data collection activities of the Center are subject to the specific provisions of both the Privacy Act of 1974 (P.L. 93-579) and the Health Services Research, Health Statistics, and Medical Libraries Act of 1974 (P.L. 93-353). The Privacy Act covers all federally sponsored and operated data collection that involves creation of a system of records containing unique personal identifiers, while the latter is limited to a portion of the Public Health Service.

In keeping with the requirements of the Privacy Act of 1974 and Center policy, each individual, household, or establishment asked to provide data to the Center is informed in writing of: (1) the authorization for soliciting the information, and that disclosure of such information is voluntary, (2) the principal purpose or purposes (usually statistical or research) for which the information is intended to be used; (3) the routine uses (as published in the "Federal Register") which may be made of the information, and (4) the effects on him, if any, of not providing all or any part of the requested information.

The Center is also bound by the provision of its authorizing legislation that the information obtained in the course of its statistical activities may not be used for any purpose other than that for which it was supplied and that such information may not be published or released in other form

if the particular establishment or person supplying the information or described in it is identifiable unless such establishment or person has consented to its publication or release in other form. Thus, the Center assurance to respondents that data confidentiality will be assured has a legal basis.

Almost all of the data that NCHS collects are requested under a pledge to the respondent, either an individual or establishment, that the information will be used only for statistical purposes and will not be released in such a way that individual persons, households, or facilities will be identifiable. The Center never releases information that would identify an individual or household. Some information is released that identifies individual facilities. When it is planned to release such information, facility administrators are informed at the time they provide the data which, if any, items will not be accorded confidential treatment.

The NCHS Policy Statement on Release of Data for Individual Elementary Units and Special Tabulations presents and explains in detail the confidentiality policy.

Data Collection Mechanisms

The Center does not have a large data collection staff of its own. It collects most of its data through interagency agreements with the Bureau of the Census or through contracts with nonfederal organizations.

It has, for a number of years, been NCHS policy to increase the capacity at the State level to collect and utilize data and to identify, as in the case of vital statistics data, instances when it is possible to unduplicate the collection of data by having it collected at the most appropriate point (Federal, State, or local), and then making it available to others who have need for the data. While vital statistics is currently the only area in which this method of data collection and sharing is completely relied upon by NCHS for the production of national data, it is hoped that in the years to come other types of data may be collected through State systems. For example, at the present time systems are being built in the areas of health manpower and facilities data and experimental work has begun in other selected areas to determine the feasibility of applying this method of cooperative data collection to the production of national statistics.

Data Release Mechanisms

NCHS releases its data in several ways. The Vital and Health Statistics series contains detailed reports on the design of the various data collection systems, findings of methodological studies, and detailed cross-tabulations of data. These reports are mailed, as issued, to individuals who have requested to be on one or more of the Center's mailing lists. An annual Current Listing and Topical Index to the Vital

All NCHS mortality and morbidity information is coded to the most recent revision of the ICD. Data for the approximately ten year period prior to December 31, 1978 are coded to the Eighth Revision, International Classification of Diseases, Adapted for Use in the United States, and data collection/reported beginning in January 1979 are coded to the ninth revision.

Data Release Mechanisms

NCHS releases its data in several ways. The Vital and Health Statistics series contains detailed reports on the design of the various data collection systems, findings of methodological studies, and detailed cross-tabulations of data. These reports are mailed, as issued, to individuals who have requested to be on one or more of the Center's mailing lists. An annual Current Listing and Topical Index to the Vital and Health Statistics Series is published that contains a cumulative list of these reports. Single copies of Center reports are available on request at no charge.

The Center also issues an annual booklet that describes public use data tapes available for purchase. This program, begun in 1969, permits researchers to perform their own analyses of the data. Tape availability and contents, along with information on ordering the tapes, are detailed in each year's Standardized Micro-Data Tape Transcripts.

The Advance Data reports are a timely but brief (4-12 pages) summary of newly available data on topics of special interest. These reports have replaced supplements to the Monthly Vital Statistics Report for all survey programs except the natality and mortality followback surveys. Both of these reports are mailed, as issued, to persons on mailing lists for the survey program from whose data the report is generated.

Many requests for unpublished data are filled regularly using tabulations that have been previously compiled.

In addition, many special tabulations are prepared each year to meet data requests that cannot be fulfilled in any other way. When special tabulations are necessary the requestor is given a cost and time estimate. Tabulations and public use data tapes are carefully reviewed to insure that confidentiality is maintained.

BASIC VITAL STATISTICS

BACKGROUND

Basic vital statistics provided through the registration system come from records of live births, deaths, fetal deaths, induced terminations of pregnancy, marriages and divorces or dissolutions of marriages. Registration of these events is a local and State function, but uniform registration practices and use of the records for national statistics have been established over the years through cooperative agreements between the States and the National Center for Health Statistics (NCHS) and its predecessor agencies.

PURPOSE AND SCOPE

The purpose of the basic vital statistics program is to formulate and maintain a cooperative and coordinated vital records and vital statistics system, promoting high standards of performance. The program is nationwide in scope covering the entire population of the United States.

DATA COLLECTION PROCEDURES

Both provisional and final vital statistics are derived from the registration system. The provisional data are obtained from counts of vital records registered without reference to the date the event occurred and the final data are obtained from the record and its contents, processed by date of occurrence of the event.

The civil laws of every State provide for a continuous and permanent birth, death, and fetal death registration system. In general, the local registrar of a town, city, county, or other geographic place collects the records of births and deaths occurring in the area, inspects, queries, and corrects if necessary, maintains a local copy, register, or index, and transmits them to the State health department. There the vital statistics office inspects the records for promptness of filing and for completeness and consistency of information; queries if necessary; numbers, indexes, and processes the statistical information for State and local use; and binds the records for permanent reference and safe keeping. Microfilm copies of the individual records or machine-readable data are transmitted to the National Center for Health Statistics for use in compiling the final annual national vital statistics volume.

The system for collecting national data on marriages and divorces is not as well developed as the system for births and deaths. All States have marriage and divorce laws but, as of January 1, 1978, three States did not have a central file of marriage records and four States did not have a central file of divorce records. Forty one (41) States and the District of Columbia had been admitted to the marriage-registration area (MRA) and

28 States to the divorce-registration area (DRA). It is these registration-area States that provide microfilm copies of their marriage and divorce certificates or machine-readable data from which NCHS derives data on characteristics of marriages and divorces and the persons involved. Even for final data, only counts of events are provided by States that have central files but are not in the registration areas. For States without central registration, final counts are collected from individual counties either by the State vital statistics office or by NCHS. In these instances some counties report only marriage licenses issued and divorce petitions filed rather than marriages performed and divorces granted.

Most States submit the microfilm copies of vital certificates or machine-readable data tapes prepared from them to the National Center for Health Statistics within 90 days following the end of each data month. Special arrangements for less frequent shipments are made with smaller registration areas. Relevant information derived from State queries of the original certificate and received by the State after shipment to NCHS are filmed and submitted with the first shipment following their receipt by the State.

Provisional vital statistics are collected and published monthly and summarized annually. They are derived from monthly reports from the States to NCHS giving the number of certificates accepted by the State for filing between two dates a month apart, without regard to actual date of occurrence. These reports to NCHS are to be mailed on or before the 25th of the month following the data month. They are the source of the provisional vital statistics published in the Monthly Vital Statistics Report (MVSR) and the Annual Summary of the MVSR. Provisional data also include a 10-percent sample of death certificates, the Current Mortality Sample (CMS), which provides provisional cause-of-death data on a monthly basis. The CMS is selected by NCHS from the regular data file of deaths for those States submitting their entire months file by the end of the following month. Otherwise the State is asked to provide a sample of records on a current basis. The sample is selected by including each record with a given last digit in the certificate number.

COLLECTION FORMS

To promote uniformity in the statistical information collected from States and local areas for national statistical purposes, the National Center for Health Statistics recommends standard certificates or reports for birth, death, fetal death, induced termination of pregnancy, marriage, and divorce. The standard certificates and reports are developed cooperatively with the States and local areas and the Federal agency, taking into account the needs and problems expressed by the major providers and users of the data. They are reviewed about every ten years to assure that they meet to the fullest extent feasible current needs as legal records and as sources of vital and health statistics. Although the use of standard certificates and reports by States is voluntary and their form and content may vary according to the laws and practices of each State, the certificates and reports in most States closely follow the standard.

DATA PROCESSING METHODS

Machine-readable data are accepted from States participating in the Cooperative Health Statistics System (CHSS). Data from all other States are coded and entered on magnetic tape.

All death and fetal death records are processed. For States not participating in CHSS, a 50-percent sample of live births is processed by selecting records with image numbers terminating in an even number. For CHSS States the entire birth file is utilized. In general, all States in the marriage and divorce registration areas that do not submit these data under CHSS send microfilm copies of all their marriage or divorce certificates to NCHS where a sample of the records is selected and processed.

RELEASE OF DATA

Vital statistics data reach the public through the Monthly Vital Statistics Report (MVSR); supplements to the MVSR; Vital Statistics of the United States (VSUS) bound volumes issued annually; micro-data tapes (issued annually); and Vital and Health Statistics Series reports and special reports. Unpublished tabulations may also be available upon request.

The MVSR's containing provisional monthly counts are published within 60 days following the end of the data month. Provisional cause-of-death information from the Current Mortality Sample are published a month later; the advance supplements to the MVSR releasing final annual statistics are published within 12 months of the data year; data tapes and unpublished tabulations are released at the same time as these advance MVSR supplements and thus, have a corresponding lag; series and special reports are issued as resources permit; and the annual VSUS's are expected to soon be distributed with about an 18-month lag.

For vital records names and addresses are not coded and never appear on data tapes or tabulations; certificate numbers are never released without written permission from the States whose records are involved; and data tapes including data from localities having less than 250,000 population are classified to reflect population size but do not reveal specific geographic areas.

USES OF THE DATA

Vital statistics data are collected, processed, and made available for public use.

Vital statistics are used in making population estimates and projections, in assessing the health of the U.S. population, in pinpointing health problems, in measuring progress made by national health programs, in epidemiological studies, in marketing research, in demographic, sociological, and economic studies, studies of the family, and in numerous and varied

other aspects of research aimed at understanding our society and its problems and progress.

Among the principal users of the data are other parts of the Public Health Service, such as the Center for Disease Control, the Food and Drug Administration, and the National Institutes of Health; other Federal agencies, such as the Bureau of the Census, Federal Reserve Board, Bureau of Labor Statistics, and Departments of Agriculture, Defense, and Transportation; members of the United States Congress; State and city governments; public and private research institutions; life insurance companies; the faculty and students of universities; newspaper reporters and feature writers; and physicians and workers in health information and education groups.

VITAL STATISTICS FOLLOWBACK SURVEYS

BACKGROUND

The National Mortality Survey and the National Natality Survey are periodic data collections based on samples of registered deaths and births occurring during a calendar year. Mortality surveys were conducted annually from 1961 through 1968, and natality surveys from 1963 through 1969 and again in 1972. A National Infant Mortality Survey was also conducted from 1964-66.

PURPOSE AND SCOPE

The national followback surveys extend for statistical purposes the range of items which are normally included on the vital records. They provide national estimates of births and deaths by characteristics not available from the vital registration system. They also serve as a basis for evaluating the quality of information reported on the vital records.

SAMPLE DESIGN

The birth or death record serves as the sampling unit, and samples of these units are selected from a frame of records representing births or deaths registered during a given period (usually a calendar year).

The sampling frame for the National Mortality Survey is the Current Mortality Sample (CMS), the 10-percent systematic sample of death certificates received each month by the National Center for Health Statistics from the registration areas in the United States. The sample for the National Mortality Survey is a sample subselected monthly from the CMS.

The sampling frame for the National Natality Survey is the file of microfilm birth certificates received each month by NCHS from the birth-registration areas of the United States. Each registration area assigns a number to each certificate prior to or during filming of the birth record. A given number of consecutive images constitutes a primary sampling unit. From each primary sampling unit, one record is chosen at random.

The sample size and the total response rate for each of the surveys was as follows:

Sample Size	Survey	Response Rate
1 out of 330 deaths	National Mortality Survey, 1961:	
	Hospitals and institutions	98%
	Informants providing names of hospitals and institutions	93%

1 out of 330 deaths	National Mortality Survey, 1964-65	91%
1 out of 110 infant deaths	National Infant Mortality Survey, 1964-65	88%
1 out of 1,000 births	National Natality Survey, 1963:	
	Mothers	86%
	Physicians	93%
	Medical facilities	98%
	Dentists	97%
1 out of 1,000 births	National Natality Survey, 1964-66	89%
1 out of 1,000 births of white infants	National Natality Survey, 1967	91%
1 out of 500 births of all other infants	National Natality Survey, 1968-69	87%
1 out of 1,000 births of white infants		
1 out of 500 births of all other infants	National Mortality Survey, 1966-68	92%
1 out of 260 of all deaths of persons 35-84 years of age		
1 out of 500 births	National Natality Survey, 1972:	
	Mothers	72%
	Physicians	72%
	Hospitals	85%

DATA COLLECTION PROCEDURES

Data for all the followback surveys are collected primarily by mail. In the natality surveys, from addresses given on the birth certificates, questionnaires are sent to the mother, the physician who delivered the baby, and the medical facility where the baby was born. For the mortality surveys, a questionnaire is sent to the person who provided the funeral director with the decedent's personal information for recording on the certificate. This questionnaire requests socioeconomic information about the decedent as well as the names and addresses of hospitals and institutions which might have provided care to the decedent at any time during his last year of life. If the death occurred in a

hospital or institution, a hospital questionnaire is sent directly to the hospital or institution asking for information about the care provided and for the names and addresses of other medical facilities providing care.

CONTENT OF THE QUESTIONNAIRES

The questionnaires for national mortality surveys have contained questions concerning the patient's last year of life. The 1961 survey included questions on hospital utilization, diagnoses, operations performed, institutions in which hospitalized, number of hospital stays, place of death, income, and whether working or retired during most of last year of life. The 1962-63 questionnaire included a few questions on hospital and institutional care in the last year of life, household composition, education, and income, with detailed questions on place of residence of the decedent. The 1964 and 1965 questionnaires included questions on health insurance, place of hospital and institutional care, the charges for hospital care and who paid them, operations performed, how much of the surgeon's bill was paid by insurance, household composition, assets, and income. The 1966-68 survey also included questions on the smoking habits of the deceased and questions regarding the household and income.

The questionnaire for the 1964-66 National Infant Mortality Survey included questions on hospitalization of the infant who died, information about other children of the mother, household composition, income, employment of mother, education of mother and father, and health insurance.

The national natality surveys collected information from mothers who had live births during a given year. The 1963 survey gathered information on the medical and dental care and radiological treatment of the mother, employment and education of mother and father, and family income. The 1964-66 survey gathered information on pregnancy history, expectations of having more children, household composition, income, whether this was a first or later marriage, date of first marriage and date of this marriage, whether mother was employed and when during her pregnancy she stopped working, education, and health insurance. The 1967-69 survey dealt with medical care and smoking habits of the mother, breast feeding, information on earlier pregnancies, expectations of having more children, whether this was first marriage, dates of first and present marriages, education and religious preference of husband and wife, information on household composition, work history, and income. The 1972 survey covered sources of medical care, health status of mother and infant, health insurance, outcomes of all pregnancies, whether pregnancy was wanted and expectations of having more children, whether the mother had an operation to prevent future pregnancies, whether this is first marriage, dates of first and present marriages, education of mother and father, and information about the household and family income.

DATA PROCESSING METHODS

After all methods of obtaining complete questionnaires have been exhausted, the data are edited, coded, and transcribed onto punch cards, with 100-percent verification. Basic range edits are made to eliminate punching errors, and the cards are then used as input for magnetic tape. Computer processing includes consistency checks, interval edits, estimation or assignment of weights, and imputation of missing data.

RELEASE OF DATA

The followback survey data have been released through the Vital and Health Statistics series reports, supplements to the Monthly Vital Statistics Report, and micro-data public use tapes, as the data became available. Confidentiality of all responses to the survey is maintained.

USES OF THE DATA

Data from the followback surveys are in demand by all major users of vital statistics data. They help meet the expanding needs for natality and mortality data in public health, medical research, epidemiology, demography, and other related fields. Data provided through the natality surveys are used by agencies and individuals responsible for or studying maternal and child health care programs, and by researchers and others studying social and demographic issues.

NATIONAL SURVEY OF FAMILY GROWTH

BACKGROUND

The National Survey of Family Growth (NSFG) is a multipurpose statistical survey that provides a wide range of information serving needs of persons and organizations concerned with the dynamics of population change, family planning, and maternal and child health. Developmental funds and necessary positions were provided for establishing the NSFG in NCHS in FY 1971, and full funding for the program was provided in the budget for FY 1972 and subsequent years. Field work for the first cycle of the survey was begun in July 1973 and completed in January 1974. The second cycle began in January 1976 and was completed in September 1976. A third cycle is expected to begin in 1980.

PURPOSE AND SCOPE

The NSFG is designed to produce data on factors influencing trends and differentials in fertility, family planning practices of the population, sources from which family planning advice and services are obtained, the effectiveness and acceptability of the various methods of family planning, and those aspects of maternal and child health that are most directly related to fertility and family planning. The survey is based on personal interviews with women in childbearing ages selected from a nationwide area probability sample of households.

SAMPLE DESIGN

The National Survey of Family Growth is based upon a cross-section sample of women in the coterminous United States, 15-44 years of age, who are or have been married, or have never been married but have children of their own living with them in the household. Excluded from the sample are women living in group quarters (i.e., five or more unrelated persons 18 years or older, unrelated to the head of the household, who live and eat together, or six or more unrelated adults who live and eat together--e.g., college dormitories, barracks, long-term care institutions), and never-married women without children. The design is a multi-stage, area-probability sample.

In the first cycle of this survey, the first stage, primary sampling units included 101 Standard Metropolitan Statistical Areas, counties, parts of counties and independent cities. Secondary sampling units consisted of enumeration districts or block groups, within selected PSU's. Where feasible, secondary units were subdivided into third-stage listing units with a probability of selection proportional to estimated housing. In the fourth stage, housing units within the third-stage listing units were chosen by systematic sampling. A fifth level of sampling among eligible women within a household was required to obtain one and only one interview in each household with at least one eligible respondent. To assure sufficient precision for presenting intragroup comparisons on various fertility variables, the black population was oversampled.

The overall sample design called for 10,000 completed interviews, including approximately 4,000 black women and 6,000 women of white and other racial groups. The response rate was an overall 81 percent. The final number of interviews was 9,797, comprised of 3,856 black respondents, and 5,941 white and other respondents.

For the second cycle, a multi-stage area-probability design based on 79 primary sampling units was developed. The remaining stages were broadly similar except that a stratum for new housing (housing built since 1970), sampled from building permit listings, was added. Overall expected sample size remained at 10,000 with comparable racial composition. After a certain level of effort was completed in the field work, remaining nonrespondents were subsampled at a rate of 1 in 2. All subsampled cases (nonresponse sample households and nonresponse eligible extended-interview respondents) were assigned a weight of 2. The response rate for Cycle II was an overall 83 percent. The actual number of completed extended interviews was 8,611, comprised of 2,946 black respondents and 5,665 white respondents.

DATA COLLECTION PROCEDURES

Data collection for the two completed cycles of the NSFG were conducted by private contractors according to NCHS specification. Personal indepth interviews were conducted with women identified as eligible, extended-interview respondents, through a household screening interview.

The data collection instruments (questionnaires) included a household screener designed to obtain household composition data and to identify eligible extended-interview respondents in the sample households. Two questionnaire versions were used for the extended interview--a currently married questionnaire for women who were married at the time of the interview, and a post-married questionnaire for women who were widowed, divorced, separated, or who had never been married but had their own children living in the household. The primary differences between the two versions were question rewording, and deletion of questions related to husbands, to make the post-married questionnaire appropriate for respondents not married at the time of the interview. Interviewer instructions were used throughout the questionnaires to skip respondents over questions, or entire questionnaire sections, which were not applicable to their individual situations.

Building upon procedures developed in Cycle I, standardized interviewer training programs were conducted in several different locations in the country which provided indepth training on the questionnaires and on the NSFG concepts and procedures. In addition to successfully completing training, interviewers were required to conduct, for review and approval,

several practice interviews prior to beginning their field assignments. Only female interviewers, and observers, were used in the NSFG. Field supervisors received the same indepth training as the interviewers, in addition to intensive training in field procedures.

Several quality control procedures were developed in Cycle I, and refined in Cycle II, to assure the quality of the collected and processed data; i.e., a systematic field edit of selected interview items (PFR/CFR) and a validation of a sample of each interviewer's work (sample recheck) were performed throughout the field work. Data preparation was validated through a 5-percent sample recode of all questionnaires in addition to systematic verification of each coder/keyer's work. A comprehensive legal code/consistency cleaning program was developed and used for the data tapes.

CONTENT OF THE QUESTIONNAIRES

The questionnaires for the NSFG cover such topical areas as marital history, a detailed pregnancy history, fecundity and expected or intended future births, pregnancy planning practices and utilization of specific contraceptive methods, the source and financing of family planning services, and a broad range of socioeconomic and demographic characteristics.

In the first cycle of the survey, the average length of an interview was 73 minutes; in the second cycle the average length of an interview was reduced to about 5[illegible] minutes.

It is anticipated that once the basic survey instrument and procedures are well established, subsamples of respondents will be selected as panels for longitudinal followup interviews. This procedure will provide prospective data and the opportunity to "verify" it in relation to actual subsequent experience.

RELEASE OF DATA

The publication program for the NSFG includes (1) the first advance report in the form of a supplement to the Monthly Vital Statistics Report (MVSR), and subsequent advance reports issued in the Advance Data series, (2) methodological reports in Series 2 and indepth substantive analyses to be presented in Series 23 of the Vital and Health Statistics Series, and (3) a public-use tape. Publications from the first cycle were issued beginning in 1976; the public-use tape for this cycle is available. This schedule is substantially longer after the close of field work than is planned for future cycles, due to the sizable amount of developmental work in the first cycle. Advance reports from the second cycle will be issued beginning in 1978; the public-use tape is expected to be available early in 1979.

USES OF THE DATA

The rapidly growing need for collection of data on trends and differentials in the birth rate, family planning, and related aspects of maternal and child health has been recognized by the President and Congress. Attention has been directed toward studying population trends, assessing the implications of future population growth, making appropriate policy recommendations, and helping to educate the Nation regarding the consequences of population growth. More information is needed than is presently available from the regular data collection programs and occasional surveys of the Bureau of the Census and the National Center for Health Statistics. The NSFG has been developed to provide, on a periodic basis, detailed data on factors influencing fertility, such as desired family size, birth spacing intentions, and family planning practices, in order to interpret current trends in the birth rate and to prepare more realistic projections of future population growth. In addition, for use in effective planning, management, and evaluation of the expanded family planning programs, new information will be provided concerning the number of couples that are unable to control their fertility to the extent that they want to, their social and economic characteristics, the nature and severity of the problems they face, and the extent to which efforts to help them are succeeding.

HEALTH INTERVIEW SURVEY

BACKGROUND

The Health Interview Survey (HIS) is the principal source of information on the health of the civilian noninstitutionalized population of the United States. The National Health Survey Act of 1956 provided for a continuing survey and special studies to secure on a non-compulsory basis accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States and the services rendered for or because of such conditions. The survey referred to in the Act, now called the Health Interview Survey, was initiated in July 1957. In its early years the survey was known to many as the National Health Survey, the name now given to a broader program of surveys in the National Center for Health Statistics.

PURPOSE AND SCOPE

The purpose of the survey is to provide national data on the incidence of illness and accidental injuries, the prevalence of diseases and impairments, the extent of disability, the utilization of health care services, and other health-related topics. The strength of this survey lies in its ability to display these variables by the demographic characteristics of the U.S. population.

The data are obtained from the people themselves and, therefore, measure the social and economic dimensions of health--the impact of illness and disability and resulting uses of health care services by individuals.

Interviews are conducted each week throughout the year in a probability sample of households. The interviewing is performed by a permanent staff of carefully trained and supervised interviewers of the Bureau of the Census under detailed specifications of the National Center for Health Statistics. Data collected over the period of a year are published as annual estimates of the health characteristics studied and form the basis for studying trends in those characteristics.

The survey covers the noninstitutionalized civilian population of the United States living at the time of the interview. Because of technical and logistical problems several segments of the population are not included in the sample or in the estimates from the survey. Persons excluded are: patients in long-term care facilities for the handicapped (data are secured on patients in some of these facilities through the Nursing Home Survey of NCHS); persons on active duty with the Armed Forces (though their dependents are included); and persons who have died during the calendar year preceding the interview. The result is that the Health Interview Survey data somewhat underestimate levels of disability and health services utilization when the total population is considered.

SAMPLE DESIGN

The sample is a multistage probability design which permits a continuous sampling of households. The first stage consists of a sample of 376 primary sampling units (PSU's) drawn from approximately 1,900 geographically defined PSU's that cover the 50 States and the District of Columbia. A PSU consists of a county, a small group of contiguous counties, or a Standard Metropolitan Statistical Area. Within PSU's, smaller units called segments are defined in such a manner that each segment contains an expected four households.

Each calendar year the sample is composed of approximately 40,000 households containing about 120,000 persons. The households in each week's sample are a probability sample representative of the target population.

The annual response rate of HIS is usually at least 95 percent of the eligible households in the sample; the 5-percent nonresponse rate is divided equally between refusals and households where no eligible respondent could be found at home after repeated calls.

DATA COLLECTION PROCEDURES

Data are collected through a personal household interview conducted by interviewers employed and trained by the Bureau of the Census according to procedures specified by the National Center for Health Statistics.

All adult members of the household 19 years of age and older who are at home at the time of the interview are invited to participate and to respond for themselves. The mother is usually the respondent for children. For individuals not at home during the interview, information is provided by a responsible family member (e.g., spouse, parent, or adult son or daughter) residing in the household. Between 65 and 70 percent of the adults aged 19 or older are self-respondents. Upon occasion a random subsample of adult household members is selected to respond to questions on selected topics. There are also instances in which followup supplements are completed for either the entire household or for individuals identified as having particular health problems. As required, these supplements are either left for the appropriate person to complete and return by mail, or the interviewer calls again in person or by telephone to secure the information directly.

Nationally there are approximately 110 interviewers, trained and directed by health survey supervisors in each of the 12 Census Bureau Regional Offices. The supervisors are career Civil Service employees whose primary responsibility is the Health Interview Survey. The interviewers are part-time employees, selected through an examination and testing process. Interviewers receive thorough training in basic interviewing procedures and in the concepts and procedures unique to the Health Interview Survey.

CONTENT OF THE QUESTIONNAIRE

The structure that has evolved for the questionnaire is one of a relatively stable nucleus of questions, approximately 70 percent of the questionnaire, complemented by one or more supplements on topics which vary from year to year.

Each year's questionnaire contains questions in the following areas:

the basic demographic characteristics of household members, including age, sex, education, and family income;

disability days, including restricted activity, bed, work and school-loss days, occurring during the two-week period prior to the week of interview;

the physician and dental visits occurring during the same two-week period;

the acute and chronic conditions responsible for these days and visits;

long-term limitation of activity and the chronic conditions related to the disability;

all hospital episodes, including the reason for entering the hospital, whether surgery was performed, and the length of stay, during the 12 months prior to interview;

the interval since the last doctor and dental visit.

Each year's questionnaire also includes a set of questions related to chronic conditions of one body system. Over a six-year period, all major body systems are covered. For example, in 1975 the chronic conditions section of the questionnaire focused on conditions of the digestive system; in 1976, on conditions of the skin and musculoskeletal system; and in 1977, on impairments. For each system, the questions are designed to obtain information on the prevalence of specific conditions and on the disability and use of health services associated with them.

The supplements to the questionnaire change each year. The 1977 topics included disability, stroke, a hearing scale for persons reporting hearing problems and a second set of health habit questions. Supplements for 1978 were on health insurance coverage, usual source of health care, blood donations, immunization, smoking, military service, and for households in the first quarter sample health expenditures in calendar year 1977. During 1979 the immunization and smoking supplements will be continued and supplements on home health care, eye care, residential mobility and retirement income will be added.

On the average the interviews require about 45 minutes in the household. Depending upon the family size and the nature and extent of its health conditions the length of interview usually runs between 15 and 90 minutes.

The content of the Health Interview Survey questionnaire reflects both the results of survey research and requests for data. For example, the recall period for many questions is the two weeks prior to the week of interview, but for some items for which the memory loss is known not to be excessive, recall extends over the year prior to interview or for some other period specified according to topic, such as three months for x-ray visits. The body-systems approach to chronic conditions was adopted in 1968, after several years in which there was instead a standard set of questions focused around a list of diverse chronic conditions. Limiting questions to one body system at a time has been found to secure more thorough reporting of conditions from a targeted list and to increase the number of conditions for which estimates of prevalence can be made.

Suggestions and requests for special supplements are received from many sources, including a panel of leading members of the health professions, university-based researchers, administrators of national organizations and programs in the private and public health sectors, and other parts of the Department of Health, Education, and Welfare (e.g., the National Institutes of Health and the Center for Disease Control). Although it is not possible to include all of the suggested topics, every effort is made to be responsive to the data needs of such groups. A lead time of at least one year is required to develop and pretest questions for new topics to be included as special supplements.

DATA PROCESSING AND QUALITY CONTROL METHODS

Throughout the data collection and processing phases there are extensive quality control activities. Each interviewer edits her completed work before returning it to her regional supervisor, and in the Regional Office there is a preliminary edit of all questionnaires. As part of the quality control program, interviewers are provided feedback on their errors. In addition, a sample of each interviewer's households is reinterviewed by a supervisor or senior interviewer for a portion of the questionnaire.

The Regional Offices forward the questionnaires to the National Center for Health Statistics for coding and data processing. At this time each questionnaire is checked again for completeness of field coverage, omissions and other errors. Illnesses, diseases and injuries reported by the household are coded to a modified version of the International Classification of Diseases. Coding for demographic and occupational items is based on that used by the Bureau of the Census in order to provide comparability. All coding is subject to recoding on a sample basis to insure a high level of accuracy.

RELEASE OF DATA

Data release occurs in several forms. The earliest reports containing survey data were issued in Series B and C, Health Statistics from the U.S. National Health Survey. About 50 reports were published in these series before they were replaced by the Vital and Health Statistics series. More than 100 reports have been published in Series 10 of the Vital and Health Statistics series. Publication of a year's data begins in about October of the year following completion of data collection. The first report, "Current Estimates," is followed by between 10 to 15 other publications.

Since not all possible cross-tabulations can be analyzed and published in Series 10 reports, many unpublished tabulations are routinely made available upon request. In addition, within budgetary and other limitations, special tabulations are prepared upon request.

The other medium for data release is public use computer tapes. The tape is usually available for distribution within about two years after completion of data collection. Primarily this time lag is due to the need for staff to work with the data in order to insure completeness and reliability of data secured on the supplements and to permit development of adequate documentation. About one month is required after receipt of a tape order to process the request and deliver the tape. Since the tape program was begun for data from the 1969 survey, approximately 45 tapes have been sold to researchers wishing to perform their own analyses from the data.

USES OF THE DATA

Data from the Health Interview Survey have been used in a number of major government programs. Estimates of health problems and resultant utilization of health care services were used in formulating the legislation for both the Medicare and Medicaid programs and for making preliminary estimates of the likely costs of the programs. Statistics on health characteristics of smokers and non-smokers provided one basis for the 1964 report of the Surgeon General entitled Smoking and Health. Special compilations on particular population groups have also been developed and provided to various White House conferences and for other special purposes. Data requests from government agencies usually account for approximately one-fourth of each year's inquiries.

Almost 40 percent of the data requests come from private industry and foundations, and an additional 10 percent are from individuals affiliated with educational institutions. In addition to general research, these users have relied upon Health Interview Survey data for estimates of persons with particular health problems for whom advertising campaigns or new products are being designed. The potential demand for special aids or devices for persons with limitations of mobility and for various types of drugs is often assessed using data from the Health Interview Survey.

Edited micro-data tapes of data from the annual survey, beginning with the 1969 data collection year, have been available for several years. Universities and Federal, State, and local governmental agencies have been the main purchasers of the tapes.

HEALTH AND NUTRITION EXAMINATION SURVEY

BACKGROUND

The first Health and Nutrition Examination Survey, referred to as the HANES, was initiated in 1970, and data collection began in April 1971. This survey, however, was an extension and expansion of the earlier Health Examination Survey (HES). The modification reflected the assignment to NCHS of an additional specific responsibility--to measure the nutritional status of the population and then to monitor changes in that status over time. The Health Examination Survey, which had been initiated a decade earlier and had carried out three separate programs, was restructured to combine the new task with its original purpose.

PURPOSE AND SCOPE

The HANES and its predecessor program share a common purpose--the collection and utilization of data which can be obtained only by direct physical examinations, clinical and laboratory tests, and related measurement procedures. This information, which cannot be furnished by the people themselves or by the health professionals who provide their medical care, is of two kinds: (a) prevalence data for specifically defined diseases or conditions of ill health; and (b) normative health-related measurement data which show distributions of the total population with respect to particular parameters such as blood pressure, visual acuity, or serum cholesterol level. The surveys use scientific samples of the U.S. population to provide representative national data which are analyzed and made available in a series of reports. Successive surveys in the HES and HANES programs have been directed to different segments of the population and have had different sets of target conditions. Thus, the first Health Examination Survey or "cycle" involved examining a sample of adults with the focus primarily on selected chronic diseases. The second and third cycles of the HES were directed respectively to children between the ages of 6 and 11 and youths between the ages of 12 and 17. Both of these surveys emphasized growth and development data and sensory defects. The nutrition component of the first HANES program was directed to a probability sample of the broad age range 1 to 74 years, while the detailed health examination component focused on the population between ages 25 and 74.

SAMPLE DESIGN

All of the HES and HANES programs have involved the use of multi-stage, highly clustered probability sampling. All of them involve collaboration with the Bureau of the Census in the development and execution of the sample design. All of them include stratification by broad geographical regions and by population density groupings. In some of them, an additional basis of stratification has been the rate of change in population between successive decennial censuses. Within the strata the subdivisions considered in the sampling process have been, in turn, the primary sampling unit, the Census enumeration district, the segment,

the household, and finally, the individual person. Using controlled selection procedures in accordance with the jointly agreed upon sample design, the Bureau of the Census selects the primary sampling units (SMSA's, counties or clusters thereof) to be included in a particular survey. The next stage of the sampling process involves selection within each of these primary sampling units of a specified number of enumeration districts, and this is followed by selection within the chosen enumeration districts of clusters of a specified number of households (presently eight) that are adjacent and make up what is designated as a segment. The sample selection process to this point is carried out centrally and makes use of various records provided by the Bureau of the Census.

The next stage of the sampling is conducted in the field in the particular chosen area. It involves interviewer visits and questionnaire completion at each one of the selected households, with the final selection of individuals included in the sample being dependent upon information in this questionnaire. Certain subgroups in the population (e.g., preschool children, the poor, etc.) presumed to be at high risk of malnutrition are oversampled at known rates. The size of the sample in the survey programs to date has varied. In each of the first three HES programs, the sample size was approximately 7,500 persons, while in the first HANES program the sample selected for the nutrition examination was about 30,000 persons, about one-fifth of whom were selected to also receive the detailed health examination.

DATA COLLECTION PROCEDURES

NCHS employees do the major part of the interviewing, history taking, examining, testing, and measuring, that provide the data. Data collection is done by specially trained teams of interviewers and examiners including physicians, nurses, dentists, dietitians, and medical laboratory and X-ray technicians. The examinations take place in the Survey's mobile examination centers. These are sets of specially constructed units, each consisting of several truck-drawn trailers which are interconnected and which provide a standardized environment for the performance of the specific parts of the examination. This is necessary for elements such as exercise testing (requiring temperature and humidity control), testing of hearing (requiring hearing chambers within which the ambient noise level conforms to A.S.A. standards for acoustical measurements), laboratory work (requiring special equipment and space where hematological tests can be made shortly after specimen collection and where blood serum and urine can be properly prepared for shipment and refrigerated prior to shipment), and the like.

Given the available personnel and dollar resources and the quantities of these resources required to collect the kinds of data obtained in these programs, it has not been possible to complete data collection in a year or even two. The general pattern has been one of a data collection period of about three years to obtain the number of

examinations required to provide valid national data. This imposes a limitation on the kinds of data to be collected by this mechanism, since conditions which might show marked year-to-year variation or seasonal patterns cannot be included. However, many important chronic diseases and health-related measurements are not subject to such changes in prevalence within short-run periods. Distributions of population according to unassisted visual acuity levels, prevalence of such conditions as diabetes or hypertensive heart disease and the like, may vary over long periods of time, but not so rapidly as to prevent data collection over a three-year period from giving a correct picture of the population levels during or at the midpoint of that period.

Voluntary sample surveys always present a problem if no data are collected on a large fraction of the selected sample because of unwillingness of individuals selected to participate in the program. In the HES and HANES programs there has always been and continues to be much attention devoted to the question of the response rate, which is the proportion of sample persons who are actually examined. In the HANES there have been, as anticipated, more problems in the area of response than had been encountered in the earlier HES programs. The difficulties faced have led to a variety of innovative measures, including a policy of remunerating examined persons. The first HANES program succeeded in obtaining household interview data on 99% of the sample population. More detailed health data appear in the medical history questionnaires, and these were completed for 88% of the selected sample persons. Finally, in this first HANES program, 74% of the sample persons selected for the nutrition component and 70% of the persons for the detailed health component were given the standard examinations and tests. We thus have considerable ancillary information on most of the not-examined persons in the sample population, and we make use of that data in the process of imputation. There is, moreover, some evidence that data obtained through examinations, tests, and measurements such as used in these surveys are less susceptible to potential bias from a given rate of nonresponse than data provided by the individuals themselves.

CONTENT OF THE QUESTIONNAIRES, TESTS, AND EXAMINATIONS

The kinds of information collected in the HANES and other examination survey programs are so varied and extensive that they are only illustrated here. With respect to nutrition, four types of data are included: (1) information concerning dietary intake--the mechanisms used have included 24-hour recall interviews and food frequency questionnaires, both administered by an interviewer who has been a trained dietitian; (2) hematological and biochemical tests--a sizable battery of such tests has been performed, at the mobile examination centers where necessary, but for the most part at a central nutrition laboratory established at the Center for Disease Control; (3) body measurements--the battery used is especially important in connection with infants, children, and youths where growth may be affected by nutritional deficiencies; and (4) various signs of high risk of nutritional deficiency, based on clinical examinations.

The health component of the HANES program includes specific defined diseases and conditions for which detailed examinations, tests, questionnaires, etc., are developed to obtain a measure of prevalence levels. These vary with the particular program and have included such conditions as chronic rheumatoid arthritis and hypertensive heart disease. Important normative health-related measurements such as height and weight are also obtained. An important element in the health component in the first HANES program was an assessment of unmet health needs through the use of index conditions. Here, for example, the examination established the presence or absence of emphysema or other chronic respiratory disease. At the same time, information was obtained from the examined person with respect to his self-perceived health needs in this regard and the actions he has taken with respect to seeking medical care. By interrelating these two kinds of information we hope to have measures of unmet health needs.

The time required for the examinations varies, of course, with the content of the examination. The time constraint introduced among the planning factors is that the total examination time not exceed 2½ hours. There is an additional time burden on the sample person arising out of the interview and completion of forms and questionnaires in the household, one that should not exceed 2 hours. The amount of time required to travel to the examination site also represents a variable burden to the sample person.

DATA PROCESSING METHODS

The data collected in the HANES program require a variety of data processing methods. X-rays must be interpreted; blood and other laboratory specimens must be processed through the appropriate laboratory operations. Certain data, such as electrocardiographic tracings, record directly onto magnetic tape and must undergo appropriate processing to translate them into digital tape form. Examination record forms and interview and questionnaire data must be coded and put onto magnetic tape. Methods for handling each specific examination element are chosen on the basis of appropriateness. The program makes use of precoded forms, marked-sense record forms, self-administered forms, interviewer-administered forms, automatic recording devices, photographs, etc. The nature of these recording methodologies gives some suggestion of the required data processing methods. Throughout the process, constant emphasis is placed on quality control measures, and various editing, verifying, and replicating steps are aimed at maximizing the reliability and validity of the resultant data. The resultant data are analyzed frequently with the collaboration of outside consultants who are expert in specific sub-specialty areas, and reports of findings are prepared.

RELEASE OF DATA

Information release from the HANES occurs primarily through publication of separate reports in NCHS Vital and Health Statistics, Series 11. Data are also released through sale of data tapes, publications in appropriate scientific journals, separate monographs, and special reports. Reports are not issued on a set frequency, but rather made available as completed. The reports are organized on a topical basis with, for example, one report presenting data on periodontal disease, another, data on auditory acuity, etc.

USES OF THE DATA

Because of the diverse nature of the data, the primary users of the data may be considered as a number of separate sets, some of which overlap with others. The reports of dental findings, for example, do not have the same audience as the findings on visual acuity or on nutrition status. Some individuals and organizations are users of most or all of the HANES and HES data, for example, policy makers and planners in Federal, State, and local health agencies. This heterogeneity of users is, of course, a major reason for the pattern of organization of the reports on a subject-matter basis. An illustration of the variety of users can be provided. The reports on body measurements and like growth and development data, for example, have been requested by large numbers of health workers in the fields of nutrition, child health, pediatrics, anthropology, medical education and health research. The American Academy of Pediatrics requested that copies be made available to its fellows. In addition to health workers, these body measurement data have been requested by many industrial groups and agencies, including automobile manufacturers. Other subject-matter areas have users of data which will include some of the same persons, but also many different ones. It might be noted that some of the primary users of HES data are the individuals who author text and reference materials used in medical schools and by health workers generally. Through this means the HES normative data reach a vastly wider audience. The sizes of the mailing lists vary with the subject, but several thousand copies are distributed for each of them. The sale of additional copies is handled by the Government Printing Office, whose sales range from hundreds to several thousand copies of each separate report.

The sale of data tapes from this program is, of course, very different in volume since the numbers of users who have the necessary computer capabilities and the need for data in full detail is much more limited. However, although the numbers of data tape purchasers are measured in dozens rather than in thousands, they include important uses. Some of these users are medical research institutions and individual researchers who wish to explore in detail specific elements of the data. A number of tapes, for example, have been sold to clothing manufacturers who wish to study interrelationships of detailed body measurements. The potential uses of data collected in the HANES are so numerous that it is impossible with existing or expected NCHS resources to fully exploit the data, and the data tape release program makes it possible for us to insure that these data are available for more detailed study.

HOSPITAL DISCHARGE SURVEY

BACKGROUND

The Hospital Discharge Survey (HDS) is the principal source of information on inpatient utilization of short-stay hospitals. Data collection began in 1964 and has been continuous since then.

PURPOSE AND SCOPE

The purpose of the HDS is to produce statistics that are representative of the experience of the U.S. civilian population discharged from short-term hospitals. Specifically, the survey provides information on the characteristics of patients, the lengths of stay, diagnoses and surgical operations, and patterns of use of care in hospitals of different size and ownership and in the four regions of the country.

The scope of HDS is limited to discharges from nonfederal hospitals in the 50 States and the District of Columbia. Only short-stay hospitals with six or more beds and an average length of stay for all patients of less than 30 days are included in the sample.

SAMPLE DESIGN

The unit of enumeration in the survey is a hospital discharge. The sample plan is basically a two-stage stratified design. The first stage is a sample of about 10 percent of the hospitals, excluding Federal hospitals, listed in the Master Facility Inventory (sampling frame). The primary stratification variables are bed size and geographic region. Hospitals are selected in direct proportion to size such that hospitals with 1,000 or more beds are selected with certainty and hospitals with less than 50 beds are sampled with a probability of approximately 1/40. Growth in the inventory of hospitals is represented in the survey by a sample of hospitals selected from a special universe of new hospitals.

The second stage of the design is a systematic sample of the discharges from the sampled hospitals. The sampling frame in nearly all hospitals is the daily listing of discharges. The size of the within-hospital sample varies inversely with the size of the hospital from about 1/100 in hospitals with 1,000 or more beds, to 4/10 in hospitals with less than 50 beds. The overall sampling rate for each bed-size group is about 1/100, the product of the first and second stage sampling rates.

In 1977 the sample consisted of 535 hospitals from a universe of approximately 7,500 short-stay hospitals. Of the 491 in-scope hospitals, information was collected from 423 participating hospitals (approximately an 86-percent response rate) on approximately 225,000 discharges.

DATA COLLECTION PROCEDURES

The Bureau of the Census, acting as the data collecting agent for NCHS, inducts sample hospitals into the HDS. After induction, hospitals are visited at least once a year by a representative of the Bureau of the Census, at which time survey procedures are reviewed and information about the hospital is updated.

Discharge data are collected throughout the year. Sample discharges are systematically selected, usually on the basis of the final digit(s) of the patient's medical record number. For each sample discharge, an abstractor records personal, administrative, diagnostic, and surgical information from the face sheet of the patient's medical record onto a Medical Abstract Form. Data collection frequency depends upon the arrangement made with the hospital. In about 35 percent of the participating hospitals, a representative of the Bureau of the Census visits the hospital bimonthly, completes the abstract forms for records selected during the previous visit, and selects records for abstracting at the next visit. This allows time for records to be completed and properly filed (or pulled from file) prior to the visit. In about 65 percent of the hospitals, the same forms are completed by members of the medical record department. All completed forms are forwarded to one of the Census Regional Offices for review and then to NCHS for coding and data processing operations.

CONTENT OF MEDICAL ABSTRACT FORM

The Medical Abstract Form contains items relating to the personal characteristics of the patient including birth date, sex, color and marital status but not name and address; administrative information including admission and discharge dates, discharge status, and medical record number; and medical information including diagnoses and surgical operations or procedures. It is estimated that medical record personnel can sample and complete each form, on the average, in about five minutes.

The contents of the Medical Abstract Form did not change from the inception of the survey until 1977 when modifications were made so that it more nearly parallels the Uniform Hospital Discharge Data Set. The items added to the abstract at that time are residence of patient (zip code), expected source of payment, disposition of patient, and dates of procedures. In 1968-1970, actual hospital charges by service and payments by source were recorded on a Ledger Abstract Form for approximately one-third of the sample discharges.

DATA PROCESSING METHODS

After transmittal from the hospital, abstract forms are subject to two reviews, two machine edits, and two quality control procedures. Forms are reviewed for completeness at the Census Regional Offices and either forwarded to NCHS or returned to the hospital for more information. Upon

receipt at NCHS, forms are again checked and, if necessary, returned to the hospital. After review, up to five diagnoses and three operations are coded according to the International Classification of Diseases as slightly modified for use with HDS. Coded information is then keyed from abstract to disk and a preliminary machine edit program checks for missing, invalid, and inconsistent codes. The information is corrected, if necessary, and then transferred from disk to computer tape.

RELEASE OF DATA

Annual data are published in the NCHS Vital and Health Statistics Series, Series 13. The publication program is, at a minimum, to update the non-medical, the medical, and the surgical data for characteristics of patients and hospitals. Special reports on average length of stay, patient charges, geographic utilization, hospital ownership, and methodology are also published.

Unpublished data are available on request from the Hospital Care Statistics Branch, which receives about 500 requests per year, usually for specific diagnostic and surgical listings in the ICDA. In addition, data for years since 1969 are available on magnetic tape.

The names of the participating hospitals and all information related to individual patients are confidential. No data are released in published, unpublished, or tape form that could identify hospitals or patients.

USES OF THE DATA

The HDS is the principal source for national data on the characteristics of patients discharged from short-stay hospitals. The data are used for a variety of planning, administrative, and evaluation activities by governmental, professional, scientific, educational and commercial institutions as well as by private citizens. The wide variety of uses of HDS data is best exemplified by the diversity of users. These include Federal agencies such as the National Institutes of Health, the Center for Disease Control, and the Indian Health Service; universities and medical schools; professional organizations such as the American Medical Association, American Hospital Association, and World Health Organization; hospitals; medical research laboratories; pharmaceutical and medical supply manufacturers; publishing houses; markets research groups; and insurance companies.

NATIONAL AMBULATORY MEDICAL CARE SURVEY

BACKGROUND

In May 1973, the National Center for Health Statistics inaugurated the National Ambulatory Medical Care Survey (NAMCS) on a continuing basis to gather and disseminate statistical data about ambulatory medical care provided by office-based physicians to the population of the United States. The need for such a system had been recognized many years before, having been pointed out as early as 1953 by the Subcommittee on National Morbidity Surveys of the U.S. National Committee on Vital and Health Statistics. During a five-year feasibility study period beginning in 1967, three major national pilot studies and several small area studies, combined with extensive consultation with experts in various areas of medical care delivery, were employed in developing the current NAMCS instruments and procedures. Through this process the current NAMCS materials and procedures have been refined to request only information considered essential to adequately describe the utilization of ambulatory services, and to require a minimal amount of time for the participating physicians.

PURPOSE AND SCOPE

The purpose of the NAMCS is to meet the needs and demands for statistical information about the provision of ambulatory medical care services in the United States. Ambulatory services are rendered in a wide variety of settings, including physicians' offices, neighborhood health centers, and hospital outpatient facilities. It is expected that the NAMCS will, in the future, encompass all of these settings, and appropriate survey instruments and methodology will be developed as resources permit. Initially, the NAMCS target population consists of all office visits within the coterminous United States made by ambulatory patients to nonfederal physicians who are in office-based practice and engaged in direct patient care. Excluded are visits to hospital-based physicians, visits to the specialists in anesthesiology, pathology, and radiology, and visits to physicians who are principally engaged in teaching, research, or administration. Telephone contacts and non-office visits are also excluded. Since about 70 percent of all direct ambulatory medical care visits occur in physicians' offices, the current NAMCS design provides data on the majority of ambulatory care services.

SAMPLE DESIGN

The most objective and reliable sources of data about physicians' services rendered to ambulatory patients during office visits are the physicians themselves and members of their office staff. The sampling frame is a list of licensed physicians in "office-based, patient care" practice compiled from files that are classified and maintained by the American Medical Association (AMA) and the American Osteopathic Association (AOA). These files are continuously updated by the AMA and the AOA, making them as current and correct as possible at the time of sample selection.

The NAMCS utilizes a modified probability-proportional-to-size sampling procedure using separate sampling frames for Standard Metropolitan Statistical Areas (SMSA's) and for nonmetropolitan counties. After sorting and stratifying by size, region, and demographic characteristics, each frame is divided into sequential zones of one million residents, and a random number is drawn to determine which primary sampling unit (PSU) from each zone is included in the sample. The NAMCS final first-stage sample contains 87 PSU's, corresponding to individual counties or small groups of contiguous counties across the country.

The second-stage sample is selected from the list of physicians located in the sample PSU's ordered by major specialty categories so that the overall probability for including any individual physician is the reciprocal of the number of physicians in the frame at the time of selection. The present annual sample consists of approximately 3,000 physicians. The sample physicians are randomly distributed across the 52 weeks of the year so that the resulting data reflect any seasonal variations. Since the assignment of the reporting week is an integral part of the sample design, each physician is required to report during his predetermined period, and no substitute reporting periods are permitted. Approximately 80 percent of the eligible physicians in the sample participate in the survey. From this size physician sample, information is secured from about 60,000 patient visits a year. Samples for subsequent years exclude with certainty sample physicians included within the previous two years of the study.

The final stage involves sampling patient visits within a physician's practice. The sampling rate, which is determined at the time of the interviewer's appointment, is dependent on the number of days during the reporting week that the physician is in practice and the number of patients he expects to see. In actual practice, the sampling procedure is handled through the use of a Patient Log. (See Data Collection)

FIELD PROCEDURES

To maximize participation levels and minimize data reporting burden in the physician's office, NAMCS field procedures have been designed to accommodate the circumstances of individual physicians. Each physician is contacted by several means, including mail, telephone, and personal interview. Initially, each physician in the sample is sent an introductory letter from the Director of the Center, followed by a letter of endorsement from the AMA or AOA. The physician is then telephoned by an informed and trained interviewer who explains the survey briefly and arranges a personal appointment to relate more detailed instructions. During this appointment, the interviewer verifies the physician's eligibility for participation in the survey, delivers survey materials with printed instructions, provides detailed verbal instructions, and assigns a predetermined seven-day (Monday through Sunday) reporting period. Also, during this appointment data concerning basic practice characteristics, such as the physician's specialty and makeup of his staff, are collected for use in analysis.

Before the beginning and again during the physician's assigned reporting week, the interviewer contacts the sample physician to answer possible questions and to ensure that procedures are going smoothly. At the end of the week, the participating physician mails his finished forms to the interviewer, who edits them for completeness before transmitting them for central data processing.

DATA COLLECTION PROCEDURES

The actual data collection for the National Ambulatory Medical Care Survey is carried out by the participating physician, aided by his office assistants when possible. The physician completes a Patient Record for a sample of his patients seen during his assigned reporting week. Based on the physician's own estimate of patients expected to visit during the survey period, the physician is assigned to use an "every-patient" or a "patient-sampling" procedure. These sampling procedures are designed so that Patient Records are completed each day of practice for at most ten patient visits. Physicians expecting 10 or fewer visits per day record data for all of them, while those expecting more than 10 visits per day record data after every second, or third, or fifth visit, observing the same predetermined sampling interval continuously. These procedures minimize the workload of data collection and maintain equal reporting levels among sample physicians regardless of the size of their practice. Each form requires one to two minutes to complete, so that approximately 15 minutes are required on days when patients are attended in the physician's office.

Two data collection forms are employed by the participating physician: the Patient Log and the Patient Record. The Patient Log is a sequential listing of patients that serves as a sampling frame to indicate for which visits data should be recorded. The Patient Record contains 14 items of information about the visit: date and duration of the visit; patient's birthdate, sex, color, and problem; whether the patient has been seen for the particular problem before and whether the patient was referred by another physician; length of time since onset of the problem; diagnoses; diagnostic and therapeutic services; seriousness of the condition; and disposition. In the future, ad hoc supplementary items will be added to the basic Patient Record to investigate specific health conditions or other aspects of ambulatory care.

DATA PROCESSING METHODS

Initial edits for completeness of the Patient Record are done by the interviewer as the forms are received from the participating physician. These forms are then sent to central processing where they are clerically edited for consistency. The two medical questions are then coded: the patient's symptoms are coded using the National Ambulatory Medical Care Survey Symptom Classification, and the physician's diagnoses are coded using the International Classification of Diseases. The data are then keypunched and put onto magnetic tape, at which time additional matching edits and consistency checks are run.

RELEASE OF DATA

NAMCS results in the form of summary statistical tabulations of national and regional estimates for number of visits, percent distributions, and population rates of use are published as soon as possible after each annual cycle is complete. More detailed tabulations and analyses follow, which present visit characteristics by major physician specialty groups, patient groups, diagnostic categories, treatment provided, and disposition. Cross tabulations of less common visit characteristics will be published when sufficient data about them are available to meet practical standards of precision. These data are released through the Center's Vital and Health Statistics Series (Series 13). Other modes of publication, including the various journals and newsletters of the NAMCS endorsing organizations, will also provide a means of releasing data to the medical community. In addition, data tapes will be made available when documentation of the data is complete and its validity is checked.

As part of the procedures designed to protect the identity of the patients, the Patient Logs containing the names of the patients are detachable from the Patient Records. At the end of the reporting period, the Patient Logs, and thus the names of the patients, are retained by the physician. Sample physicians are further assured that "all information which would permit identification of an individual, a practice or an establishment will be held confidential, will be used only by persons engaged in and for the purposes of the survey, and will not be disclosed or released to other persons or used for any other purpose." All findings are released in the form of summary statistics which preclude any individual identification.

USES OF THE DATA

NAMCS provides a wide range of baseline data on the characteristics of the users and providers of ambulatory medical care. These baseline data, together with trend data as the survey progresses, will provide new insights into ambulatory medical care and stimulate further research on the use, organization, and delivery of ambulatory care.

The manner in which ambulatory care is actually dispensed, as described by the survey, can be related to contemporary arrangements for educating and training physicians and other health personnel to provide ambulatory medical care, particularly primary care. Such comparisons will suggest useful new directions and emphasis that can be incorporated in the curricula, helping to make medical education more responsive to the shifting requirements and health problems of the population. NAMCS information will be useful to health planning agencies, managers of health delivery systems, and others concerned with planning, monitoring, and managing health care resources, and it will be valuable to those developing and evaluating new and modified health care systems and arrangements. The continuing nature of the survey will permit observation and measurement over time of different methods for managing and treating patients'

problems, and it will provide general information on the etiology and epidemiology of selected conditions. It will also provide valuable information about the speed and effectiveness with which certain advances in medical practice are adopted by the basic source of medical care, the office-based physician.

NATIONAL NURSING HOME SURVEY

BACKGROUND

Between 1963 and 1969, the National Center for Health Statistics conducted surveys of nursing homes and their residents on an ad hoc basis. With the implementation of the Medicaid and Medicare programs, the increased utilization of nursing homes, and the projected increases in the aged population, those who set standards for, plan, provide, and assess long-term care services needed comprehensive national data on a continuing basis. To meet their needs, the National Nursing Home Survey System was developed in 1972, with the initial survey in the System conducted in 1973-74, the second conducted in 1977, and future surveys planned every three years.

PURPOSE AND SCOPE

This continuing data collection system is a series of nationwide sample surveys of nursing homes, their residents and staff. The purposes of the surveys are:

- To collect national baseline data on characteristics of the nursing home, its services, residents, and staff for all nursing homes in the Nation, regardless of whether or not they are participating in Federal programs such as Medicare or Medicaid.

- To collect data on the costs incurred by the facility for providing care by major components such as labor, fixed, operating, and miscellaneous costs.

- To collect data on Medicare and Medicaid certification (such as utilization of certified beds and the health of residents receiving program benefits) so that all data can be analyzed by certification status.

- To provide comparable data for valid trend analyses on a variety of topics. Such analyses can, for example, identify the impact of legislative changes in standards and reimbursement on the growth of facilities and the impact of institutionalization on the health of the aged.

- To interrelate facility, staff, and resident data to reveal the relationships that exist between utilization, services offered, charges for care, and the cost of providing care.

For the initial survey conducted in 1973-74, the universe included only those nursing homes which provided some level of nursing care, regardless of whether or not they were participating in the Medicare or Medicaid programs. Thus, homes providing only personal or domiciliary care were excluded. Beginning with the 1977 survey, the universe was expanded to include all nursing, personal care, and domiciliary care homes,

regardless of their participation in Medicare or Medicaid. Homes which provide room and board only are excluded. In both surveys, homes in the universe included those which were operated under proprietary, nonprofit, and government auspices. The universe included homes which were units of a larger institution (usually a hospital or retirement center).

SAMPLE DESIGN

The Master Facility Inventory (MFI) listing is the universe from which the sample homes are selected. The MFI listing, maintained by NCHS, contains basic information about the home (such as name, address, size, ownership, number of residents, and number of staff) that is needed to design efficient sampling plans.

Resident data are collected by reviewing medical records and questioning the nurse who usually provides care for the resident. Residents are not interviewed directly. The response rates for the surveys differed according to the type of questionnaire, as presented below.

Questionnaire	Response Rate	
	1973-74 Survey	1977 Survey
Facility	97%	95%
Expense	88%	85%
Staff	82%	81%
Resident	98%	99%
Discharges	-	97%

The initial survey, conducted from August 1973-April 1974, had a nationally representative sample of 2,100 nursing homes, with a subsample of 25,000 staff and 20,000 residents. The second survey, conducted from May-December 1977, had a total sample of 1,700 nursing homes, with a subsample of 18,900 staff, 8,000 residents, and 5,900 discharged residents.

DATA COLLECTION PROCEDURES AND QUESTIONNAIRE CONTENT

The survey system uses several questionnaires. The facility questionnaire includes questions on number of beds and residents, services provided, certification status, and various utilization measures. The expense questionnaire includes questions on the facility's expenses by major components, such as labor, fixed, operating, and miscellaneous expenses. The staff questionnaire includes questions on training, previous experience, salary, duties performed, and fringe benefits. The current-resident questionnaire includes questions about the resident's demographic characteristics, health status, functional status, participation in social activities, monthly charge, and source of payment. Included in the 1977 survey was the discharged-resident questionnaire which included some of the same questions as the current-resident questionnaire, selected on the basis of their availability in the medical record.

The survey has a number of respondents in a home and is a combination of personal interview and self-administered questionnaires. Facility information is secured through a 20-minute personal interview with the administrator. Expense data are collected on a self-administered questionnaire, requiring about 30 minutes to answer, completed by the facility's accountant under authorization from the administrator. Sampled staff members complete a brief form that requires about five minutes to complete. Information on sample current residents is secured by the interviewer in a personal interview with the nurse who provides care to the resident and who refers to information from the medical record. About 15 minutes is required for each sample resident. For the 1977 survey, information on the sample discharged residents was secured by the interviewer in a personal interview with the nurse who was most familiar with the medical records and who referred to them for replying to all questions.

The 1977 survey included several modifications of the 1973-74 survey design and methodology:

- Collecting data on discharged residents, especially in the areas of health status, length of stay, and where the resident goes after discharge.

- Collecting data on the revenues of the facility as well as on its expenses.

- Producing estimates for five States with the largest proportion of nursing home residents.

- Limiting respondent reporting by limiting subsamples of residents and discharges to a maximum of eight per facility, and of staff to 23 per facility.

DATA PROCESSING METHODS

All data items are edited to assure that all responses are accurate, logical, and complete. Those items left unanswered are imputed by assigning a value from a responding unit with major characteristics identical to those of the nonresponding unit.

RELEASE OF DATA

National estimates are available to requesters in the form of tabulations, NCHS publications, and data tapes. Tables are aggregated so that no one respondent can be identified. Similarly, no identifying information concerning the facility, staff, or residents is contained on the data tapes, although a pseudo identification number allows linkage of data among the various files.

A report analyzing provisional tabulations is available six months after data collection ends. Selected final tabulations are released six months later. Data are analyzed and presented in the NCHS Vital and Health Statistics Series, Series 13. These reports present data on the facility (utilization, expenses, services), residents (demographic characteristics, health status, services received, charges), and the staff (number, training, experience). Once data have been reviewed and analyzed, data tapes are available for release to requesters.

USES OF THE DATA

Nursing homes are the most rapidly expanding sector of the health care industry. Federal and State funds spent for nursing home care have been increasing each year. Thus, national data on costs and the characteristics of the nursing home, its residents, and staff are needed on a continuing basis by the Department of Health, Education, and Welfare for program planning and the setting of national policies. Congress needs information on the interrelationship among facility-resident-cost data in order to draft legislation on standards and reimbursement. In addition, these data are needed by professionals who conduct epidemiological research and those who plan, provide, and assess long-term care services at State and local levels.

NATIONAL REPORTING SYSTEM FOR FAMILY PLANNING SERVICES

BACKGROUND

The National Reporting System for Family Planning Services (NRSFPS) provides information about the characteristics of persons receiving medical family planning services at family planning service sites throughout the United States and its territories. Responsibility for operating the NRSFPS was delegated to NCHS in 1968 by the Office of the Assistant Secretary for Health and Scientific Affairs of the Department of Health, Education, and Welfare, which had in turn been designated by the then Bureau of the Budget as the focal point for Federal family planning statistical activities.

PURPOSE AND SCOPE

The NRSFPS was established in order to produce uniform, comparable statistical information on federally supported family planning programs and, to the extent possible, on those of other public and private organizations. Thus, included in the scope of the reporting system are all grant recipients under the Health Services Administration (HSA) umbrella (i.e., Title V, Title X, Indian Health, Community Health Center projects, etc.) as well as family planning services provided by service sites affiliated with Planned Parenthood/World Population and other private organizations that do not receive funds from any Federal agency.

SAMPLE DESIGN

From its initiation in 1968 until June 30, 1977, the system collected information from all identified family planning service sites on a 100-percent basis. As of July 1, 1977 the system was changed to a sample survey. This change was made primarily because of the large increase in the volume of records (from less than one million in 1972 to about six million in 1976) and the data processing costs associated with the increase.

The sample survey is based on a two-stage stratified sample design. For selection of the initial sample, an unduplicated list of service sites was compiled from several sources: (a) sites already enrolled in the NRSFPS; (b) an update of the 1974 Inventory of Family Planning Clinics; (c) a list of grantees under Titles V and X; and (d) a current list of Planned Parenthood/World Population affiliated clinics. Prior to sample selection, the identified universe was ordered by State groups, formed by combining States containing similar numbers of clinics. The clinics were stratified into three size classes; then, within the three size classes, the service sites were ordered by State and county.

The final 1977 first stage sample consisted of 1,500 service sites which were selected from the total universe of the approximately 6,000 family planning service sites identified in the United States and its territories. These 1,500 service sites are expected to provide information on approximately 300,000 visits per year.

The second stage sample consists of a sample of family planning visits made to the sample service sites--the sampling rate being dependent upon the size and location of the service site. Thus, the visits are sampled at a variable rate, depending on the volume of visits made to a service site. For example, some sites provide information on every patient, some sample visits on an every-other-patient basis, some on every tenth patient, etc. Nationally, the sample contains about one-fourth of the clinics and about 5.0 percent of all family planning patient visits.

DATA COLLECTION PROCEDURES

A form, the Clinic Visit Record (CVR), containing the items of information to be reported for each visit in the NRSFPS sample, is provided to the clinic for completion at the time of the patient visit. Demographic and socioeconomic characteristics (e.g., age, race/ethnicity, family size) are filled in by administrative staff from existing records and by questioning the patient. All information on pregnancy and contraceptive history, medical services provided at the visit, and contraceptive method at the end of the visit are completed by the medical provider.

For sites that provide data on only a sample of patient visits, a patient log is kept to identify the patients for whom a CVR is to be completed. At the option of the site a CVR may be completed for all patients, in which case the sample is selected either by the local data processor or by NCHS. Sample sites either may use the CVR provided by NCHS, or they may use their own forms. If a site wishes to use a form of its own, it must be sent to NCHS for review and approval prior to its use.

For sample sites which are part of a local automated system, their completed CVR's are sent to the local data processing organization where the data are coded, edited, and put on magnetic tape. The tapes from the local ADP system are then sent to the central processor, the National Center for Health Statistics. Local systems may elect to do their own sampling prior to submission of tapes, in which case the data tape will contain only the appropriate number of sample visits. However, most of the sampling will be conducted by NCHS.

CONTENT OF THE CLINIC VISIT RECORD (CVR)

The CVR contains items of information for both the sampled site and the patient. Each patient is classified as a "new" or "continuation" family planning patient. Demographic and socioeconomic information reported are date of birth, race, ethnicity, years of school completed, current student status, gross family income for the past 12 months, and the number of persons supported by that income.

For women only, a pregnancy history is obtained that includes the number of live births and the number of pregnancies that ended in stillbirth or abortion. A contraceptive use history is included for all patients to identify those who have ever used a method of birth control regularly. For those who have, there is additional information about the most

recently used method(s)--whether it is still used, how long it was used and where the method was prescribed or obtained. The contraceptive method(s) adopted at the end of the visit is indicated, as is the reason (e.g., pregnant or an infertility patient) if there is no method adopted.

DATA PROCESSING AND QUALITY CONTROL METHODS

As a part of the quality control procedure a special subsample of 200 clinics was selected in order to assess the problems encountered by site personnel in following specifications for selecting sample visits and completing the CVR. During the first year of the survey, visits to these sample clinics are planned in order to evaluate the accuracy of information being collected and to determine how well the survey procedures are being carried out. As problems are discovered, all clinics in the total sample will be contacted in an attempt to correct the problems.

Coding of all information on CVR's received in hard copy form is verified on a 100-percent basis. All data, including records received on punched cards and magnetic tape, are edited on a sample basis to determine if the records are in scope for the survey based on the definitions of a user, patient, or visit. Appropriate actions are initiated to resolve all errors, inconsistencies, and omissions of data as soon as they are detected. Further, clinics that are delinquent in reporting are contacted to determine the reason(s) for the delinquency.

RELEASE OF DATA

Annual summary estimates will be made for the Nation, the 10 DHEW regions, and each State. Preliminary estimates are expected to be available within six months of the completion of data collection. Data will be published in Advance Data Reports and in Series 13 of the Vital and Health Statistics Series.

USES OF THE DATA

A primary use of data from the NRSFPS is the annual update of the "Five-Year Plan for Family Planning Services and Population Research," a report required by P.L. 94-63 from the Secretary of the Department of Health, Education, and Welfare to the Congress. In addition, the NRSFPS provides statistics on the number of family planning visits according to patient characteristics, and on the number and types of services provided to family planning patients. The NRSFPS is a major source of data on family planning service site utilization which can be used for measuring the impact and effectiveness of the family planning program on the target population (high risk and medically indigent groups). Data obtained on users of family planning service sites can be used for various health, population, and research evaluation purposes. The data may also be used by management specialists in the field of family planning as an aid in evaluation of present family planning programs and as a tool in planning

for expanded programs. Finally, the data from the NRSFPS can be used to determine statistical relationships among the clinic users, the services provided, and the characteristics of the family planning service sites.

MASTER FACILITY INVENTORY

BACKGROUND

The Master Facility Inventory (MFI) is a comprehensive file of the 33,000 facilities in the United States which provide medical, nursing, personal, or custodial care to groups of unrelated persons on an "inpatient" (at least overnight) basis. Facilities in the MFI are categorized into three broad types: hospitals, both short- and long-stay; nursing and related care homes; and other custodial or remedial care facilities including homes or resident schools for the deaf, blind, mentally retarded, emotionally disturbed, other neurologically impaired, or physically handicapped; resident treatment centers for alcohol and drug abusers; orphanages or homes for dependent children; and homes for unwed mothers. It is the most comprehensive file of inpatient health facilities available in the United States.

PURPOSE AND SCOPE

The MFI has two basic purposes. It is an important national source of statistics on the number, type and geographic distribution of inpatient facilities in the United States. In addition, it serves as the universe from which probability samples are selected for conducting sample surveys.

DATA COLLECTION PROCEDURES

The MFI was first assembled in 1962-63, by collating the files of four Federal agencies which contained the names and addresses of facilities, directories of national associations and organizations, and State licensure files.

Two mechanisms are used to keep the data in the MFI as current as possible. For all facilities except hospitals, NCHS conducts a series of mail surveys to (1) insure that the data on file on the basic characteristics of the facilities are accurate, and (2) identify and then delete those facilities that have gone out of business or are no longer eligible for inclusion. These surveys are conducted on an approximately biennial basis.

In addition, at regular intervals State licensure agencies, national voluntary associations, and other appropriate sources send to NCHS their most recent directories or lists of new facilities. These lists are then clerically matched with the most current MFI file and facilities not already included are added.

For hospitals, data were gathered annually in a joint survey of the American Hospital Association (AHA) and NCHS. The contractual arrangement which in effect merged the hospital portion of the MFI with the AHA's annual survey of hospitals began in 1968. Through 1975, the AHA performed the data collection (during October-January each year) for its member hospitals, while NCHS performed the data collection for the approximately 400 non-AHA registered hospitals. The AHA processed both

portions of the survey according to the same edit and processing specifications and delivered to NCHS two edited tapes, one for AHA hospitals and one for non-member hospitals. Beginning in 1976 the AHA surveys all hospitals as part of its annual survey, and NCHS purchases a data tape from the AHA. All of these surveys of facilities have consistently achieved response rates of 90 percent or higher.

With the implementation of the health facilities component of the Cooperative Health Statistics System (CHSS), health facilities data for a State are collected and processed by the State CHSS contractor as a byproduct of the licensure renewal process. Edited data are then provided on computer tape to NCHS on an annual basis. Data for all facilities except hospitals in States not currently participating in the CHSS continue to be collected by NCHS on a biennial basis.

CONTENT OF DATA COLLECTION INSTRUMENTS

The following types of data are collected for the three categories of facilities:

Hospitals: Ownership; major type of service offered; whether various facilities and services are offered; number of beds, admissions, inpatient days of care, and discharges; patient census; number of bassinets, live births, and newborn days of care; outpatient uitlization; number of surgical operations; revenue, expenses, and assets; and staffing.

Nursing Homes: Ownership; major type of service; licensed and staffed beds; beds certified for Medicare and Medicaid, admission policy with regard to age, sex, and various conditions; patient census by age and sex; inpatient days of care; number of admissions, discharges, and deaths; staffing; who is in charge of nursing care; number of patients receiving nursing care; services routinely provided; basic monthly charge; and operating expenses.

Other Facilities: Ownership; major type of service; licensed and staffed beds; beds certified as intermediate care beds; admission policy regarding age and sex; patient census by age and sex; inpatient days of care; number of admissions, discharges, and deaths; staffing; basic monthly charge; and operating expenses.

Approximately 30 minutes is needed to complete one of the questionnaires.

DATA PROCESSING AND QUALITY CONTROL

The MFI is considered to be the most comprehensive file of inpatient health facilities available in the United States. In order to measure statistically the extent of its scope of coverage, a Complement Survey was developed. It is an application of a general technique often called "multiframe survey." In this application there are two frames--the Master Facility Inventory and a geographic area sample list. From a

probability area sample, all institutions found in the sample areas are identified and the probability with which each comes into the sample is determined. Those inpatient health facilities found in the area sample survey (the Complement Survey) are matched against the MFI list of facilities for that area. Any in-scope facility discovered in this sample survey but missed by the MFI constitutes undercoverage and an appropriate weight is assigned to the missed facility. Since 1973 the Current Population Survey of the Bureau of the Census has been used for the Complement Survey.

Data for nonresponding facilities, as well as data for questions not answered on the returned questionnaires, are obtained through various types of imputation procedures. Whenever possible, data from the previous MFI survey are used to replace missing data. For those responding facilities that omitted data which are not available from a previous survey, an imputation method is used. In this method all facilities of the same type are stratified by predetermined variables such as bed size groups, ownership categories, and major type of service. Missing data for a facility are then supplied by using the data from a facility of like characteristics. When a responding facility fails to answer any of the major questions (e.g., bed size, ownership, type of facility, number of patients), a followup questionnaire containing the questions they omitted is sent to them and they are asked again to answer.

RELEASE OF DATA

Data from the MFI surveys are published in various NCHS reports. The annual editions of Health Resources Statistics contain the most current available data (often preliminary estimates) from these surveys. Final data are published in Series 14 of Vital and Health Statistics and in a special series reporting data by county and SMSA (one for hospitals and one for nursing homes). Directories of facilities for the mentally retarded and of nursing homes have been published from the MFI survey data. Methodology reports appear in Series 1 and 2 of Vital and Health Statistics. In addition, the AHA's annual publication, Hospitals: JAHA Guide Issue, is based on the hospital survey.

The MFI data also are released in the form of (1) special tabulations prepared specifically for individual requests; (2) computer tapes; and (3) printouts or listings from the data tapes.

Approximately 45-50 tape copies are purchased annually by such organizations as drug companies, State governments, other Federal agencies, private contractors working on projects for other government agencies, universities, and national health care organizations.

Only the financial data from the annual hospital surveys are confidential; all other information is available for individual hospitals. Portions of the nursing home and other health facility data are, however, collected under a pledge of confidentiality and can only be released in statistical aggregates, or with all identification items stripped.

NATIONAL INVENTORY OF FAMILY PLANNING CLINICS

BACKGROUND AND PURPOSE

The data generated by this inventory serve three important purposes. First, the inventory provides characteristics data on "all" family planning clinics in the United States, regardless of their funding, and, as such, constitutes the only source for these data on a national basis. Second, these data greatly enhance the analytical potential of the National Reporting System for Family Planning Services (NRSFPS) by determining the actual universe of which the NRSFPS clinics are a part; these data, when combined with the NRSFPS data, are of great use in the evaluation of and planning for family planning facilities, as well as in research studies related to the family planning area. The third purpose of this inventory is to serve as a universe source for the NRSFPS, and for other sample surveys the Center may conduct in this area.

ESTABLISHMENT OF THE INVENTORY UNIVERSE

During 1972-73, the Center began developing the list of family planning clinics for this inventory. The first step involved contacting all possible sources that might have listings, directories, magnetic tapes, etc., of family planning clinics. These included Federal, State, and local government agencies, national organizations, and private agencies.

To determine how many long-term care facilities provided family planning services, a question was included in the Center's 1973 Master Facility Inventory (MFI) survey of all custodial and remedial care facilities. In addition, from the annual hospital surveys of the MFI, information was obtained on those hospitals that offered family planning services.

The Office of Education's 1972-73 Higher Education Directory was used as the source for a letter mailing in the spring of 1973 to all college and university health units or infirmaries to determine which ones provided family planning services on campus.

The listings, directories, forms, etc., received from these sources were merged and unduplicated, to yield a total initial list of 14,524 potential service sites.

As a final effort to "clean" the listing prior to the first national survey, a letter was mailed in October 1973 to all the listed projects to:

(1) confirm the address listings to minimize the problem of undeliverable questionnaires,

(2) confirm that each of the projects or programs did indeed provide family planning services, and

(3) determine the name and actual location of all family planning clinics for which the addressee was operationally responsible.

These mailings produced a 95-percent response rate and resulted in the reduction of the universe to 10,321 facilities, 5,857 of which provided medical family planning services and 4,464 of which were treated as nonmedical providers.

BIENNIAL SURVEY PROCEDURES

The first survey of all facilities identified in the inventory was begun in April 1974. Data collection extended from April through August 1974 and consisted of an initial mailout of the questionnaire to the 10,321 presumed providers, two followup mailings, and final telephone followups to nonrespondents. Information was obtained on the clinic's physical location; ownership; "catchment" area; funding sources; frequency and types of sessions; patient characteristics; contraceptive, medical, and ancillary services offered; and staffing. Approximately 30-45 minutes were needed to complete the questionnaire.

As a result of the processing of the survey data, those facilities that proved to be nonproviders, out of business, out of scope, duplicates, or postmaster returns, were deleted and some additions were made to the list of medical providers, resulting in a revised universe of 9,781 facilities, of which 5,719 were medical and 4,062 were nonmedical service sites. A total of 8,170 service sites responded in whole or part to this first survey, for an overall response rate of 83.5 percent.

The 1975-76 inventory contained 7,234 sites, 4,660 of which were medical and 2,574 of which were nonmedical service sites; this represented an 82-percent response rate. A 1978 inventory was conducted in the period May through August. It used a briefer questionnaire that required only about 15 minutes to complete.

USES OF THE DATA

Three Center reports deal with data from this survey. A Series 1 of the Vital and Health Statistics Series report provides a detailed description of the methodology used in assembling the inventory and in developing the survey procedures and materials. Reports describing facility characteristics and services offered are published in Series 14. In addition, a Monthly Vital Statistics Report (MVSR) supplement reported on patient and staff characteristics from the first inventory.

(1) Within the Center, the Division of Health Resources Utilization Statistics uses these data to augment their analyses of the patient data generated by the National Reporting System for Family Planning Services.

(2) The Office of Family Planning in the Bureau of Community Health Services, Health Services Administration, in conjunction with Planned Parenthood-World Population, uses these data in their preparation of the Department's annual Congressional Five-Year Update report on family planning services.

In addition to the published reports, these data will be released in the form of: (1) special tabulations prepared specifically for individual requests, (2) computer tapes, and (3) printouts or listings from the data tapes.

As these data are collected without any reference to their confidentiality they are free of such restrictions, and all data items collected are available for individual clinics.

Two mechanisms are being developed to increase the reliability of the survey data from the inventory. The first of these concerns the development of a complement universe (similar to that used for the MFI) to be used as a statistical measure of the extent of the coverage of the inventory. The second is the development of a systematic reporting mechanism to identify newly established clinics for addition to the inventory.

HEALTH MANPOWER INVENTORIES AND SURVEYS

BACKGROUND

From 1966 to 1974 the primary responsibility for the collection of health manpower data rested with the Bureau of Health Manpower (BHM). During that period the NCHS health manpower data collection activities were limited to a few allied health occupations and the NCHS published statistics on health manpower were derived largely from secondary sources. In November 1974, the responsibility for the collection of general-purpose health manpower data was transferred from BHM to NCHS. In 1977 NCHS was given the delegation of authority, under Section 708 of Public Law 94-484 (Health Professions Educational Assistance Act), for establishing a program, including a uniform health professions data reporting system, to collect, compile, and analyze data on health professions personnel.

PURPOSE AND SCOPE

The primary purpose of the Center's health manpower data collection activities is to produce reliable general-purpose statistics on the distribution and availability and the demographic, educational, and practice characteristics of persons employed in the health occupations. For the 13 health occupations that are licensed in all of the 50 States and the District of Columbia, data are collected on an inventory basis covering all licensees in each State. These occupations are: doctors of medicine, doctors of osteopathy, pharmacists, chiropractors, dentists, dental hygienists, registered nurses, licensed practical nurses, nursing home administrators, optometrists, physical therapists, podiatrists, and veterinarians. Data are obtained on the remaining health occupations largely through special-purpose surveys of facilities or other employers. This approach is required because of the large number of employers of allied health workers, the limited use of licensure in the field, the large and rapidly growing numbers and types of allied health workers, and the inadequate historical data base.

DATA COLLECTION PROCEDURES

The health manpower inventories are conducted through the mechanism of the Cooperative Health Statistics System (CHSS). Those States which are participants in the CHSS collect and process the data according to specifications provided by NCHS and provide NCHS with a data tape. For those States not currently participating in the CHSS, data for selected occupations are collected by either a private agency or an organization representing the health profession, under contract to the Center.

The data collection is performed in conjunction with each State's licensure renewal mechanism, either on an annual or biennial basis as dictated by the licensing procedures in the State. The licensure mechanism offers a number of distinct advantages for the development of a data system for the production of timely, accurate, and complete statistics on health manpower. Licensure data provide accurate counts

and distributions by geographic areas of licensed health manpower and contain the necessary information for contacting all persons employed in a health profession in order to obtain additional detail.

The inventory data are obtained by means of a self-administered questionnaire which is either mailed out as an attachment to the application for license or enclosed as a separate form. Both mail and telephone followup procedures are used with nonrespondents. Approximately 10 minutes is required for the completion of a questionnaire. Item completion rates of at least 90 percent are generally achieved and coverage approaches 100 percent of the universe.

The health manpower surveys conducted by the Center are of two types: (1) censuses of facilities to gather headcount data for detailed lists of health occupations employed in such facilities as hospitals, nursing homes, group practices and clinics; and (2) surveys of facilities to gather detailed information about specific types of health workers, such as nursing personnel in hospitals, health care administration personnel, and public health personnel.

The mail questionnaire represents the primary method of data collection for most of these surveys. Both mail and telephone followup procedures are utilized. Time required for completion of the questionnaires varies, depending upon the content of the particular survey. Response to these surveys generally exceeds 80 percent.

CONTENT OF THE QUESTIONNAIRES

For the health manpower inventories, occupation-specific minimum data sets (MDS) have been established for some of the 13 occupations. Each MDS consists of standard core items basic to all 13 health professions and additional data items unique to the specific occupation. The following items are generally included for all occupations:

Administrative--licensing State, occupation, type of licensure action, name and mailing address of licensee.

Education and training--location of school of graduation, year of graduation, education or highest degree.

Demographic characteristics--date of birth, sex, marital status, race/ethnicity.

Employment characteristics--years active in the occupation, weeks worked in the past 12 months, current activity status, geographic location of place of practice one year ago, geographic location of current primary place of work, principal setting of primary activity, number of hours usually spent per week by type of activity, and occupational specialty.

The Center's health manpower surveys vary in content and purpose. At a minimum, the censuses of facilities obtain information on full or part-time status and credentialing status. The sample surveys generally are designed to obtain data on personal characteristics, educational background, practice setting, location, and specialty.

DATA PROCESSING PROCEDURES

The data from the inventories and surveys are generally processed by the agency or organization responsible for the data collection. Standardized coding, editing, and imputation procedures are developed by the Center and provided to the contractor for purposes of insuring both the quality and comparability of the data. The Center unduplicates the edited data in order to account for persons licensed by more than one State. In addition, the Center performs considerable verification and cross-checking of the data against other lists and sources of health manpower data.

RELEASE OF DATA

The unduplicated data are summarized and disseminated by NCHS to all participating agencies in a manner that is comparable among State and local areas in terms of content, definitions, and format. The data appear in the following Center publications: Vital and Health Statistics Series 14 reports, Health Resources Statistics, and a series of State reports. Special tabulations are available upon request, and copies of many of the data tapes may be purchased. The data tapes are stripped of all identifiable information pertaining to individuals.

USES OF THE DATA

Analysis of the data is intended to fulfill the following purposes: to assess current health manpower capabilities relative to the needs for health services at the local, State, and national levels; to project future demands on the supply of health manpower; to increase the efficiency of such data collection efforts; and to reduce the effort and costs involved in disseminating and using health manpower information. The inventories can also be used as a sample frame for special surveys designed to obtain more detailed or specialized information.

U.S. GOVERNMENT PRINTING OFFICE 1979 -622-